

$$p(s, x, t, y) \geq 0, \quad \int p(s, x, t, y) dy = 1$$

Chapman-Kolmogorov:  $p(s, x, t, y) = \int p(s, x, \tau, z) p(\tau, z, t, y) dz$

$$[T_{s,t} f](x) \triangleq \int p(s, x, t, y) f(y) dy$$

$$f(y) = \mathbb{1}_B(y), \quad T_{s,t} f(x) = \int p(s, x, t, y) \mathbb{1}_B(y) dy = P_r(s, x, t, B)$$

$$T_{s,t} = T_{s,\tau} T_{\tau,t}$$

$$\begin{aligned} \int p(s, x, t, y) f(y) dy &= \int \int p(s, x, \tau, z) \underbrace{p(\tau, z, t, y) f(y)}_{dy} dz \\ &= \int p(s, x, \tau, z) T_{\tau,t} f(z) dz \\ &= T_{s,\tau} T_{\tau,t} f(x) \end{aligned}$$

$$L_t = \lim_{h \downarrow 0} \frac{T_{t,t+h} - T_{t,t}}{h} = \text{id}$$

generator of  
Markov process

$$L_t^* P = - \nabla \cdot (p \vec{b}) + \nabla^2 \cdot (\frac{1}{2} p A)$$

$$L_t f(x) = \frac{1}{2} A(t, x) \cdot \nabla^2 f(x) + \vec{b}(t, x) \cdot \nabla f(x) \leftarrow$$

$$(1) \lim_{h \downarrow 0} p(t, x, t+h, \vec{B}_{\frac{\varepsilon}{h}}(x)) = 0$$

$$(2) \lim_{h \downarrow 0} \int (\vec{y} - \vec{x}) p(t, x, t+h, y) dy = \vec{b}(t, x) \quad \odot_x$$

$$(3) \lim_{h \downarrow 0} \int (\vec{y} - \vec{x})(\vec{y} - \vec{x})^T p(t, x, t+h, y) dy = A(t, x) \quad \hat{=} 0$$

$$(1) \Pr(\|X_{t+h} - X_t\| \geq \varepsilon \mid X_t) \rightarrow 0 \text{ as } h \downarrow 0$$

$$(2) \frac{E[X_{t+h} - X_t \mid X_t = x]}{h} \rightarrow \vec{b}(t, x)$$

$$(3) \frac{E[(X_{t+h} - X_t)(X_{t+h} - X_t)^T \mid X_t = x]}{h} \rightarrow A(t, x)$$

$$\langle L_t f; P \rangle = \langle f; L_t^* P \rangle$$

$$\rightarrow \langle \frac{1}{2} A \cdot \nabla^2 f + \vec{b} \cdot \nabla f; P \rangle = \langle \nabla f; \frac{1}{2} p A \rangle + \langle \nabla f; p \vec{b} \rangle$$

$$\int f' \cdot g dx = g f \Big| - \int g' f dx$$

$$\langle f'; g \rangle = - \langle f; g' \rangle$$

$$\underbrace{\nabla \cdot \vec{pb}}_{\text{div}(p\vec{b})} = \begin{pmatrix} \partial_{y_1} \\ \partial_{y_2} \\ \vdots \\ \partial_{y_d} \end{pmatrix} \cdot \begin{pmatrix} pb_1 \\ pb_2 \\ \vdots \\ pb_d \end{pmatrix} = \sum_{i=1}^d \partial_{y_i} (pb_i)$$

$$\nabla p = \begin{pmatrix} \partial_1 p \\ \partial_2 p \\ \vdots \\ \partial_d p \end{pmatrix}$$

$$L_t = \lim_{h \downarrow 0} \frac{T_{t,t+h} - T_{t,t}}{h}$$

FPK:  $\partial_t p(s, x, t, y) = L_t^* p(s, x, t, y)$  (forward)

$-\partial_s p(s, x, t, y) = L_s p(s, x, t, y)$  (backward)

Derivation:  $\langle f, \partial_t p(s, x, t, \cdot) \rangle$

$$\partial_t \langle \underbrace{f(y)}_{\parallel}; \underbrace{p(s, x, t, y)}_{\parallel} \rangle = \partial_t T_{s,t} f(x) = \lim_{h \downarrow 0} \frac{T_{s,t+h} - T_{s,t}}{h} f(x)$$

$$= \lim_{h \downarrow 0} T_{s,t} \frac{T_{t,t+h} - T_{t,t}}{h} f(x) = T_{s,t} L_t f(x)$$

$$= \langle L_t f, p(s, x, t, \cdot) \rangle = \langle f; L_t^* p(s, x, t, \cdot) \rangle$$

$$\frac{T_{s-h,s} - T_{s,s}}{h}$$

FPK for SDE: 
$$dX_t = \vec{b}_t(X_t) dt + G_t(X_t) dB_t$$

$$L_t f = \vec{b}_t \cdot \nabla f + \frac{1}{2} [G_t G_t^T] \cdot \nabla^2 f$$

$$\partial_t P(s, x, t, y) = L_t^* P = \boxed{-\nabla \cdot (P \vec{b}_t) + \frac{1}{2} \nabla^2 \cdot (P G_t^T G_t)}$$

Derivation: take a function  $f$

$$df(X_t) = \nabla f(X_t) dX_t + \frac{1}{2} \nabla^2 f(X_t) \cdot G_t G_t^T dt$$

$$= \left[ \vec{b}_t \cdot \nabla f + \frac{1}{2} \nabla^2 f \cdot G_t G_t^T \right] dt + \underbrace{\nabla f \cdot G_t}_{\text{martingale}} \cdot dB_t$$

$$\underbrace{d \langle f, P(s, x, t, \cdot) \rangle}_{\text{Ito}} = \langle \underbrace{L_t^* f}_{\text{II}} , P(s, x, t, \cdot) \rangle dt \quad \underbrace{\text{III}}_0$$

Continuity equation

$$dX_t = \vec{b}_t(X_t) dt$$

$$\partial_t P = -\nabla \cdot (P \vec{b})$$

Heat equation:

$$\partial_t P = \Delta P = \sum_{i=1}^d \partial_i^2 P$$

$$\int_0^t dX_t = \int_0^t \sqrt{2} dB_t = \sqrt{2} B_t \sim N(0, 2t)$$

$$P(t, x) \geq P(t, \cdot)$$

$$P: t \mapsto P(t, \cdot)$$

$$\frac{dP_t}{dt} = -\nabla f(P_t) = \Delta P_t$$

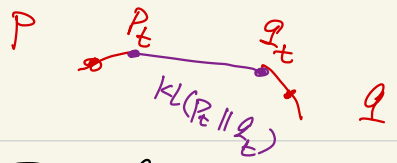
$$f(P) = \frac{1}{2} \|\nabla P\|_2^2 = \int \langle \nabla P(x), \nabla P(x) \rangle dx$$

$$\frac{d f(P + \varepsilon \cdot q)}{d\varepsilon} \Big|_{\varepsilon=0} = \frac{d \frac{1}{2} \|\nabla(P + \varepsilon q)\|_2^2}{d\varepsilon}$$

$$= \frac{d \frac{1}{2} \int \langle \nabla P + \varepsilon \nabla q, \nabla P + \varepsilon \nabla q \rangle}{d\varepsilon}$$

$$= \int \langle \nabla P, \nabla q \rangle dx = \langle \nabla P; \nabla q \rangle$$

$$= -\langle \Delta P; q \rangle$$



score matching: two densities  $P, Q$

Fisher divergence:  $F(P \parallel Q) = E \left\| \underbrace{\nabla \log P(x)}_{\vec{S}_P(x)} - \underbrace{\nabla \log Q(x)}_{\vec{S}_Q(x)} \right\|^2$

$P_t = P * N(0, 2t), \quad Q_t = Q * N(0, 2t)$

$\left. \frac{d}{dt} KL(P_t \parallel Q_t) \right|_{t=0} = - F(P \parallel Q) \left. \right|_{t=0}$

$\int P_t \log \frac{P_t}{Q_t} dx$

$\partial_t P = \Delta P \leftarrow N(0, 2t)$

$\partial_t [P * N(0, 2t)]$

$\Delta (P * N(0, 2t))$

$\int \frac{dP_t}{dt} \cdot \log \frac{P_t}{Q_t} dx + \int \cancel{P_t} \cdot \frac{\partial P_t}{\partial t} dx - \int P_t \frac{\partial Q_t}{\partial t} dx$

$\int \Delta P_t \cdot \log P_t dx - \int \Delta P_t \log Q_t dx - \int P_t \frac{\Delta Q_t}{Q_t} dx$

$\langle \Delta P_t, \log P_t \rangle = \int P_t \langle \vec{S}_P, \vec{S}_Q \rangle - \langle \frac{P_t}{Q_t}, \Delta Q_t \rangle$

$-\langle \frac{\nabla P_t}{P_t}, \nabla \log P_t \rangle = - \int P_t \|\vec{S}_P\|^2 dx = \langle \nabla \frac{P_t}{Q_t}, \nabla Q_t \rangle$

$$- \int P_t \|\vec{S}_{P_t}\|_2^2 + \int P_t \langle \vec{S}_{P_t}, \vec{S}_{Q_t} \rangle + \int \nabla \left( \frac{P_t}{Q_t} \right) \cdot \nabla Q_t$$

$$\int \frac{\nabla P_t \cdot Q_t - P_t \cdot \nabla Q_t}{Q_t^2} \cdot \nabla Q_t$$

$$\int \frac{\nabla P_t}{P_t} \cdot \frac{\nabla Q_t}{Q_t} \cdot P_t - P_t \|\vec{S}_{Q_t}\|^2$$

$$= \int P_t \cdot \langle \vec{S}_{P_t}, \vec{S}_{Q_t} \rangle - P_t \|\vec{S}_{Q_t}\|^2$$

$$- \int P_t \|\vec{S}_{P_t}\|_2^2 + \int P_t \cdot 2 \langle \vec{S}_{P_t}, \vec{S}_{Q_t} \rangle - \int P_t \|\vec{S}_{Q_t}\|_2^2$$

$$= - \int P_t \cdot \|\vec{S}_{P_t} - \vec{S}_{Q_t}\|_2^2 dx$$

$$= - F(P_t \| Q_t)$$

$N(0, D)$  data

$P, q$  find  $T$  s.t.  $q = T_{\#}^{-1} P$

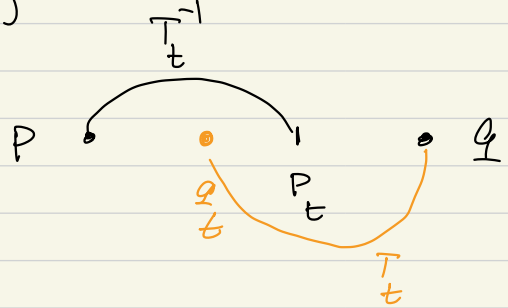
$x \sim P, T^{-1}(x) \sim q$

also  $P = T_{\#} q$

build  $T$  through  $\{T_t\}$

$P_t := (T_t^{-1})_{\#} P$

$q_t := (T_t)_{\#} q$



$P_t(x) = P(T_t^{-1}x) \cdot |\det T_t'(x)|$

$P_t(x) = P(z) \cdot \frac{dx}{dz}$   
 $\uparrow \quad \uparrow$   
 $T_t^{-1}x \quad T_t^{-1}x$

$\min_{T_t} KL(q \parallel P_t) = \int q \log \frac{q}{P_t}$

$\equiv - \int q \cdot \log P_t$

$KL(P_t \parallel q) = \int P_t \log P_t - P_t \log q$

$\approx - \sum_{i=1}^n \log P_t(x_i)$